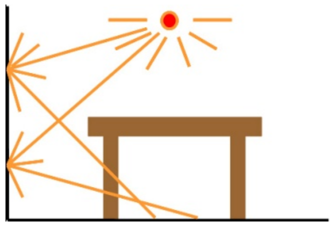
12. Globale Beleuchtung

Radiosity

Das *Radiosity*-Verfahren stammt ursprünglich aus der Wärmelehre und modelliert die Lichtausbreitung unter Beachtung des Energiegleichgewichtes in einem geschlossenen System. Das Verfahren beschreibt den physikalischen Vorgang der Ausbreitung von Licht in einer diffus reflektierenden Umgebung, also die Berechnung der Helligkeiten aller Flächen einer Szene unter Berücksichtigung der gegenseitigen Beeinflussung. Somit werden auch Flächen, die nicht direkt beleuchtet sind, eine gewisse Helligkeit erhalten. Jeder beleuchtete Gegenstand wirkt als sekundäre Lichtquelle. Für die Bildgenerierung wird zunächst die Lichtausbreitung im Raum berechnet, ohne dass die Kameraposition bekannt ist, wobei vereinfachend angenommen wird, dass der Beobachter die Ausbreitung des Lichtes nicht beeinflusst. Die Objekte können dann aus verschiedenen Richtungen dargestellt werden, ohne dass die Lichtausbreitung jedes Mal neu berechnet werden muss.



Die Radiosity Gleichung

Die Szene bestehe aus n ebenen Polygonen P_i, die beim Radiosity-Verfahren als *Patches* bezeichnet werden. Vereinfachend wird angenommen, dass jedes Patch eine homogene, perfekt diffuse Oberfläche hat. Lichtquellen sind ebenfalls Patches, die ihr erzeugtes Licht in alle Richtungen gleichmäßig abstrahlen. Die ***Radiosity*** B_i von P_i ist die gesamte abgestrahlte Energie, das ist die Summe aus Eigenstrahlung und Reflexion als Leistung pro Flächeneinheit. Diese Lichtenergiedichte ist proportional zur wahrgenommenen Helligkeit. Als nächste Vereinfachung wird angenommen, dass die Radiosity für alle Positionen auf einem Patch den gleichen Wert hat. Unter diesen Voraussetzungen sieht die Formel für die Radiosity eines Patches so aus:

$$B_i = E_i + \rho_i \cdot \sum_{j=1}^{n} B_j \cdot F_{ij}$$

Hier ist E_i die Eigenemission von P_i, ρ_i ist der diffuse Reflexionskoeffizient der Oberfläche (gibt an, wie viele % des einfallenden Lichtes diffus reflektiert wird, auch *Albedo* genannt), n ist die Anzahl der Patches der Szene, B_j sind die Radiosities aller anderen Patches, und F_{ij} sind die sogenannten *Formfaktoren*, die angeben, welcher Anteil der auf P_i wirkenden Radiosity von P_j stammt (das ist das gleiche wie der Anteil der Radiosity von P_i die auf P_j trifft). F_{ij} sind rein geometrische Größen und unabhängig von Lichtquellen oder Radiositywerten, und können daher vor der Radiositybestimmung berechnet werden.

Die Radiosity-Formeln für n Patches ergeben ein lineares Gleichungssystem von n Gleichungen in n Unbekannten B_i, das sich mittels Gauß-Seidel-Iteration numerisch lösen lässt:

$$B_i - \rho_i \sum_{j \neq i} B_j F_{ij} = E_i$$

$$\begin{bmatrix} 1 & -\rho_1 F_{12} & \cdots & -\rho_1 F_{1n} \\ -\rho_2 F_{21} & 1 & \cdots & -\rho_2 F_{2n} \\ \vdots & \vdots & & \vdots \\ -\rho_n F_{n1} & -\rho_n F_{n2} & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{bmatrix}$$

$$B_i^{k+1} = E_i + \rho_i \sum_{j \neq i} B_j^k F_{ij}$$

Dieses Gleichungssystem hat folgende Eigenschaften:

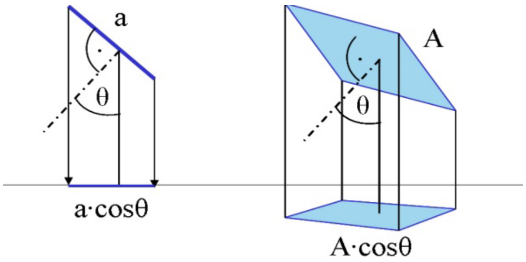
(1) $F_{ii} = 0$ für alle i, da sich ein Patch nicht selbst beleuchten kann,
(2) $\Sigma(j=1,n)F_{ij}=1$, da sich die Anteile aller Patches auf 100% ergänzen müssen,
(3) $\rho_i < 1$ für alle Patches, da nicht mehr reflektiert werden kann als eintrifft.

Daher ist die Matrix diagonaldominant und verhält sich numerisch gutartig.

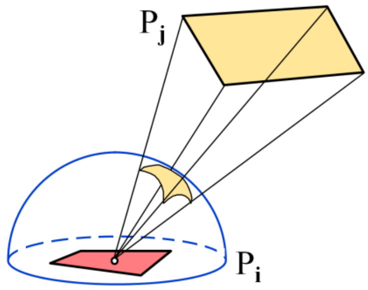
Außerdem gilt:

(4) E_i ist für die meisten Patches 0, weil normalerweise nur einzelne Patches Lichtquellen sind.

Berechnung der Formfaktoren

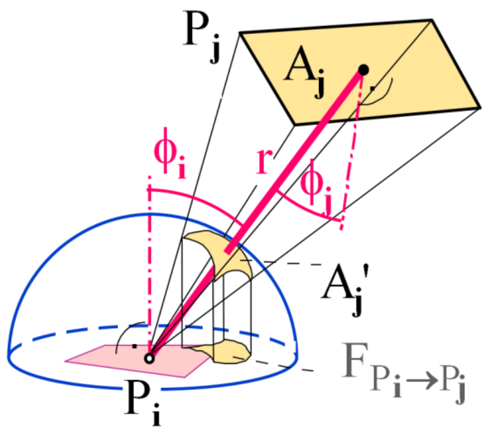


Um eine Formel für die Formfaktoren F_{ij} ableiten zu können, brauchen wir einen einfachen geometrischen Zusammenhang: Die Fläche der Normalprojektion einer Fläche A auf eine andere Fläche verkleinert sich um den Cosinus des Winkels zwischen den Flächen, ist also $A\cdot\cos\theta$.



Der *Formfaktor F_{ij}* ist nun definiert als der Anteil der vom Patch P_i ausgehenden Strahlungsenergie, die den Patch P_j trifft. Mit anderen Worten, wie viele Prozent der Energie von Patch P_i trifft auf Patch P_j. Man kann sich leicht überlegen, dass dieser Wert auch aussagt, wie viele Prozent der auf Patch P_i eingehenden Energie von Patch P_j kommt.

F_{ij} lässt sich nun so berechnen: Wir nehmen einmal an, dass die Größe der Patches klein im Verhältnis zu ihrem Abstand r ist. Sei A_j die Fläche von P_j . Über dem Patch P_i denken wir uns eine Halbkugel („hemi-sphere“) mit Radius 1, auf die der Patch P_j projiziert wird. Die daraus resultierende Fläche A_j' hat dann etwa die Größe $A_j\cos\varphi_j$, wobei φ_j der Winkel zwischen der Patchnormalen und der Verbindung zwischen den beiden Patches ist. Wenn man nun bedenkt, dass Energie, die unter einem Winkel auf Patch P_i einfällt, proportional zum Cosinus des Einfallswinkels wirksam wird, so muss man diesen Wert noch mit $\cos\varphi_i$, multiplizieren (das entspricht einer Projektion auf die Bodenfläche der Halbkugel), um den korrekten Anteil des Einflusses des Patches P_j zu erhalten. Da die Summe aller Formfaktoren F_{ij} für einen Patch P_i natürlich 1 sein muss (100%), normieren wir das Ergebnis noch mit der Größe der Bodenfläche der Halbkugel, also mit $1^2\pi = \pi$, und erhalten den Formfaktor:



$$F_{ij} = \frac{\cos\phi_i \cos\phi_j A_j}{\pi r^2}$$

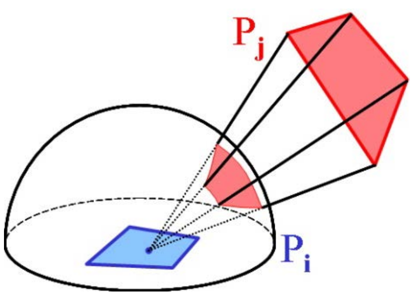
Genau genommen ist der Formfaktor nun die Summe aller Einflüsse von P_j gemittelt über die Fläche von P_i, dann ergibt sich:

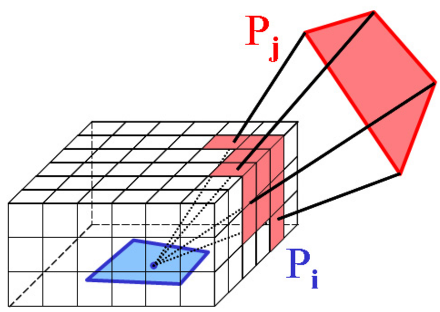
$$F_{ij} = \frac{1}{A_i}\int_{A_i}\int_{A_j} \frac{\cos\phi_i \cos\phi_j}{\pi r^2}\, dA_j\, dA_i$$

Diese Formeln gelten unter der Voraussetzung, dass sich keine Hindernisse zwischen den beiden Patches befinden, dass das Licht also ungehindert von P_i nach P_j kann. Korrekte Formfaktoren müssen also noch die gegenseitige Sichtbarkeit mitberücksichtigen.

Für die Formfaktoren gilt auch noch das *Reziprozitätsprinzip*, das die Abhängigkeit der Formfaktoren zwischen zwei Patches zueinander in Beziehung setzt: $\mathbf{A_i \cdot F_{ij} = A_j \cdot F_{ji}}$.

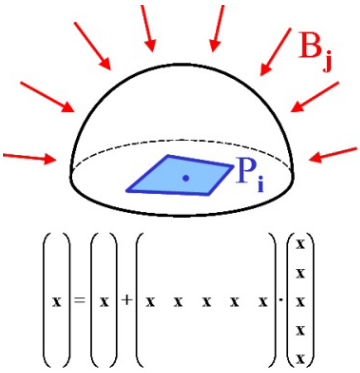
In der Praxis verwendet man zur Abschätzung der Formfaktoren statt einer Halbkugel über P_i einen Halbwürfel („hemicube“), und bildet auf diesen die gesamte Szene ab. Dazu kann man die z-Puffer-Technologie verwenden, d.h. die Oberfläche des Halbwürfels wird in regelmäßiger Weise in Pixel aufgeteilt und alle anderen Patches werden auf diese mit dem Würfelmittelpunkt als Projektionszentrum abgebildet. Für jedes Pixel bestimmt man im Vorhinein seinen Formfaktor und zählt dann für jedes Patch diese Anteile zu dessen Formfaktor zusammen. Alternativ kann man Formfaktoren auch durch Ray-Tracing-Verfahren berechnen.



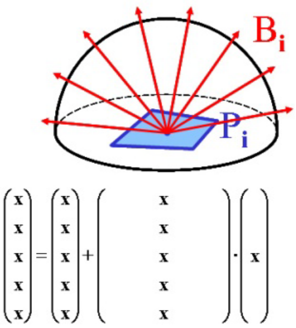


Fortschreitende Verfeinerung

Um das Gleichungssystem mit dem Gauß-Seidel-Verfahren lösen zu können, muss man alle Formfaktoren vorweg berechnen, weil man alle Einträge der Koeffizientenmatrix benötigt. Bei n Patches sind das fast n² Zahlen, was nicht nur zeitlich sehr aufwändig sein kann, sondern vor allem auch enormen Speicherplatz benötigen kann. Zur Lösung linearer Gleichungssysteme mit den erwähnten Eigenschaften eignet sich auch das Southwell-Verfahren, das man sich geometrisch so vorstellen kann: Statt in einem Schritt für ein Patch P_i die nächste Iteration B_i^{k+1} bestmöglich zu berechnen, indem dieses alle Energie aller anderen Patches einsammelt („gathering“, links) wählt man das hellste Patch aus und verteilt dessen Energie („shooting“, rechts) auf alle anderen. Dadurch werden in einem Schritt alle Patches etwas besser. Und weil man immer die Energie des hellsten Patches als nächstes verteilt, konvergiert das Verfahren viel schneller.



„gathering“



„shooting“

$$B_i^{k+1} = E_i + \rho_i \sum_{j \neq i} B_j^k F_{ij}$$

Sei $B_{(i \text{ von } Bj)}$ der Radiosity-Anteil von P_i, der von B_j verursacht wird. Der Einfluss von P_i auf B_j ist symmetrisch zum Einfluss von P_j auf B_i:

$B_{(i \text{ von } Bj)} = \rho_i B_j F_{ij}$, also $B_{(j \text{ von } Bi)} = \rho_j B_i F_{ji}$.

Daraus und aus $A_i \cdot F_{ij} = A_j \cdot F_{ji}$ folgt
$B_{(j \text{ von } Bi)} = \rho_j B_i F_{ij} (A_i / A_j)$,

es lässt sich also der Formfaktor F_{ij} benutzen. Für jeden Patch speichert man neben der bisher gesammelten Radiosity B_j (also dem bis jetzt besten Schätzwert) auch die „noch nicht verschossene Radiosity“ ΔB_j, die die Basis für die Auswahl des nächsten „hellsten“ Patches ist. Am Anfang werden die B_j und die ΔB_j natürlich mit E_j initialisiert: $B_j = \Delta B_j = E_j$ für alle j.

Vereinfacht sieht ein Iterationsschritt daher so aus:

```
select patch i with highest Aᵢ*ΔBᵢ
FOR selected patch i {set up hemicube
                      calculate form factors Fᵢⱼ   }

FOR each patch j { Δrad := ρⱼ*ΔBᵢ*Fᵢⱼ*Aᵢ/Aⱼ
                   ΔBⱼ := ΔBⱼ + Δrad
                   Bⱼ := Bⱼ + Δrad     }
ΔBᵢ := 0
```

Dieses Verfahren wird auch schrittweise Verfeinerung genannt („progressive refinement“).



3 Beispiele für Radiosity-Bilder
(© Lischinski; Hrbek; Feldman & Wallace)

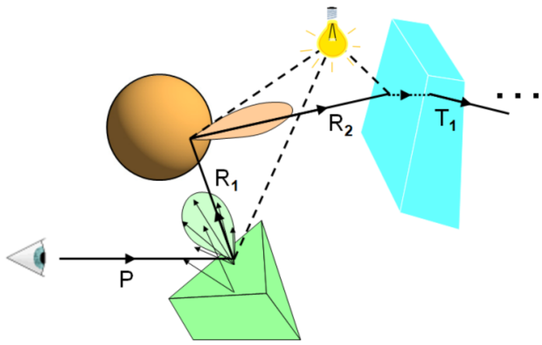
Aspekte von Radiosity

Radiosity ist eine *blickpunktunabhängige* Methode zur Berechnung der Helligkeit der einzelnen (diffusen) Patches, nach der noch ein Renderingschritt notwendig ist. Meist verwendet man ein einfaches Polygon-Verfahren mit Gouraud-Schattierung. Um im fertigen Bild auch die mit Ray-Tracing möglichen Effekte erzielen zu können, kann man die so erzielten diffusen Schattierungswerte aber auch als Basiswerte für ein Ray-Tracing verwenden. Dadurch lassen sich dann auch Spiegelungen und Schatten usw. schön darstellen.

Das hier vorgestellte Grundprinzip von Radiosity lässt sich noch in vielfacher Hinsicht erweitern. Um die Anzahl der Patches zu verringern kann man diese hierarchisch strukturieren, so dass weiter entfernte Patches nicht einzeln behandelt werden müssen. Um an Stellen, an denen sich die Beleuchtung abrupt ändert (z.B. Schattenkanten) keine falsche Verschmierung der Beleuchtung durch zu grobe Patches zu generieren, und um umgekehrt an diesen Stellen die Patches nicht zu klein werden zu lassen, verwendet man *Discontinuity-Meshing*.

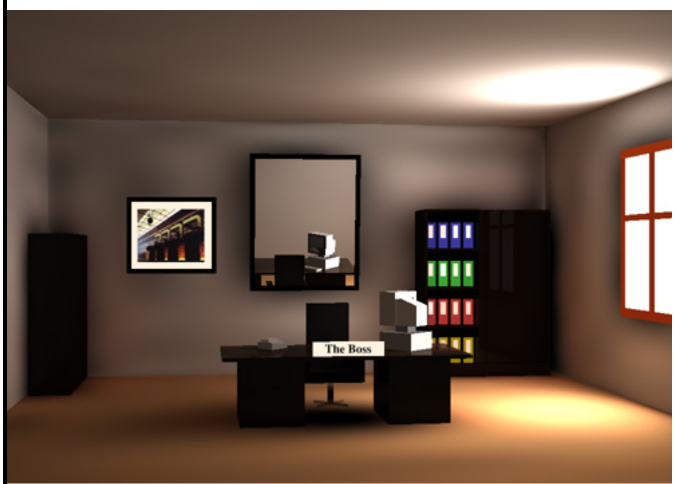
Path Tracing

Path-Tracing ist eigentlich eine Erweiterung von Ray-Tracing (wird auch „Monte Carlo Ray-Tracing“ genannt), bei der an jedem Auftreffpunkt nicht in alle relevanten Richtungen Sekundärstrahlen gelegt werden, sondern entsprechend der dort gültigen Verteilungsfunktion zufällig nur eine Richtung ausgewählt wird. Dies ermöglicht die Inkludierung von Gegebenheiten, wo viele verschiedene Lichtrichtungen relevant sind, wie diffuse Reflexion und ausgedehnte Lichtquellen. Natürlich müssen nun pro Pixel viele Strahlen berechnet und gemittelt werden, um kein zu starkes Rauschen im Ergebnisbild zu erhalten. Grundsätzlich entspricht diese Vorgehensweise der Monte Carlo-Integration eines mehrdimensionalen Integrals, dass die Lichtausbreitung im Raum beschreibt („Rendering Equation“). Die Verwendung von Quasi-Zufallszahlen statt Pseudo-Zufallszahlen reduziert dabei die Varianz merklich.



Photon Mapping

Bei der *Photon-Mapping*-Methode werden Lichtstrahlen von den Lichtquellen aus verfolgt, also ähnlich wie bei Ray-Tracing, jedoch in Vorwärtsrichtung. An Auftreffpunkten wird die Wirkung des Lichtes gespeichert und später wird zwischen diesen Werten das Aussehen des Objektes interpoliert. Dies ermöglicht es, den Effekt von Lichtquellen auch in komplizierten Situationen korrekt zu berechnen (z.B. Spiegelungen der Lichtquelle(n) oder Kaustiken). Eine Kombination von Path-Tracing und Photon-Mapping kann fast alle Lichteffekte in ein Bild integrieren.



Beispiele für Global Illumination Resultate:
links: Radiosity + Ray-Tracing (siehe Spiegelungen)
rechts: Path-Tracing + Photon-Mapping
(© Feda; VRVis)